Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 23 (2017) 624–644

Transportation Research Procedia

www.elsevier.com/locate/procedia

22nd International Symposium on Transportation and Traffic Theory

An optimization modeling of coordinated traffic signal control based on the variational theory and its stochastic extension

Kentaro Wada [a,*], Kento Usui [b], Tsubasa Takigawa [c], Masao Kuwahara [d]

[a] *Institute of Industrial Science, The University of Tokyo, Tokyo, Japan*
[b] *Nagoya Railroad Co., Ltd, Nagoya, Aichi, Japan*
[c] *East Nippon Expressway Co., Ltd., Saku, Nagano, Japan*
[d] *Graduate School of Information Sciences, Tohoku University, Sendai, Miyagi, Japan*

Abstract

This study considers an optimal coordinated traffic signal control under both deterministic and stochastic demands. We first present a new mixed integer linear programming (MILP) for the deterministic signal optimization wherein traffic flow is modeled based on the variational theory and the constraints on a signal control pattern are linearly formulated. The resulting MILP has a clear network structure and requires fewer binary variables and constraints as compared with those in the existing formulations. We then extend the problem so as to treat the stochastic fluctuations in traffic demand. We here develop an accurate and efficient approximation method of expected delays and a solution method for the stochastic version of the signal optimization by exploiting the network structure of the problem. Using a set of proposed methods, we finally examine the optimal control parameters for deterministic and stochastic coordinated signal controls and discuss their characteristics.

Peer review under responsibility of the scientific committee of the 22nd International Symposium on Transportation and Traffic Theory.

Keywords: coordinated signal control; kinematic wave theory; variational theory; random arrival; Clark approximation; cross-entropy method

1. Introduction

A system of traffic signals plays a vital role in determining the performance of arterial roads and/or road networks. Considerable research has been devoted to designing signal control parameters, such as cycle lengths, green splits, and offsets for coordinating traffic signals; several simulation-based heuristic optimization methods (e.g., Park et al., 2000; Maher et al., 2013) and commercial softwares (e.g., TRANSYT) are available for determining control parameters considering detailed and realistic phenomena and constraints (see Papageorgiou et al., 2003; Cantarella et al., 2015).

By contrast, the theoretical results for coordinated traffic signal controls are limited. Koshi (1975, 1989) showed the fundamental relationship between cycle length and delay of a signalized road under the simplification assumptions. A similar but more general result was recently reported by Jin and Yu (2015) for a stationary homogeneous ring road based on an explicit formula for a macroscopic fundamental diagram. Using a similar approach, Daganzo and Lehe (2016) explored optimal traffic coordination schemes (i.e., offsets) for different traffic density levels. However, these

* Corresponding author. Tel.: +81-3-5452-6098 (ext. 58171); Fax: +81-3-5452-6420.
E-mail addresses: wadaken@iis.u-tokyo.ac.jp (Kentaro Wada)., kuwahara@plan.civil.tohoku.ac.jp (Masao Kuwahara).

Peer review under responsibility of the scientific committee of the 22nd International Symposium on Transportation and Traffic Theory.
10.1016/j.trpro.2017.05.035

theoretical studies can only examine one signal control parameter at a time, and the relationships between the different optimal control parameters are not completely understood.

To fulfill the gap between the simulation and theoretical approaches, developing mathematical programming models capable of providing an (exact) optimal solution is necessary; this is because the solution serves as a performance bound in addition to providing insight into the fundamental characteristics of coordinated signal controls. In this direction, the suitable modeling of (i) traffic flow along a signalized arterial road and (ii) various signal constraints within an optimization framework are essential.

To address the first issue, the kinematic wave (KW) model has historically been employed to describe traffic dynamics along a signalized arterial road with several intersections. Lo (1999, 2001), the more noticeable studies, formulated a coordinated signal control problem as a mixed integer linear programming (MILP) using the cell transmission model (CTM). Han et al. (2016a) also formulated a similar control problem based on the variational theory (VT) of the KW model (Daganzo, 2005a,b). Owing to the fact that the VT approach allows link-based modeling and is free from unintended vehicle holding, the number of binary variables in MILP using CTM can be substantially reduced. However, the VT-based model does not consider the stochasticity in the traffic flow. This is in sharp contrast to the isolated traffic signal optimization wherein expected delay formulae have been established (Cheng et al., 2016). Residual queues due to stochastic arrivals could largely affect delays, particularly under the near saturation condition; thus, the optimality of the resulting control parameters by a deterministic model may not be guaranteed.

The second issue regarding signal constraints is as important as traffic flow modeling because the ways of formulating the signal constraints strongly affects the complexity of the signal control optimization. Within MILP, Lo (2001) used common control parameters (e.g., green splits, offsets) as the control variables and directly modeled the relationships between the parameters. Although this approach is straightforward, the resulting formulation is quite complex because a number of *if-then rules* need to be linearized by introducing additional binary variables. In addition, the lost times were not considered despite their importance. Lin and Wang (2004) treated the lost times as penalty terms of the objective function. However, this approach failed to take into account the queueing phenomenon at traffic signals.

This study first proposes a simpler mathematical programming formulation for deterministic coordinated traffic signal control. In our model, traffic dynamics is modeled using the variational theory of KW model, and the constraints on a signal control pattern are described as a set of constraints on a certain network. The resulting MILP has a clear network structure and the requires fewer binary variables and constraints as compared with those in existing formulations. As an extension of the basic problem, we introduce stochastic boundary conditions into the VT. By exploiting the network structure of the problem, we develop an accurate and efficient approximation method of expected delays as well as a solution method for the stochastic version of the signal optimization. Using a set of proposed methods, we finally examine the optimal control parameters for deterministic and stochastic coordinated signal controls and discuss their characteristics.

The remainder of this paper is organized as follows. In Section 2, we formulate a deterministic coordinated signal control problem based on the VT. To formulate the signal constraints with minimal number of binary variables, the network representation of these constraints is presented here. In Section 3, we extend the basic problem so as to treat the stochastic fluctuations in traffic demand, and construct an accurate and efficient approximation method for the expected delays. A solution method for the stochastic version of the coordinated signal control is also developed. In Section 4, we finally examine the optimal control parameters for coordinated signal controls under both deterministic and stochastic demands using a set of methods developed in the previous sections, and discuss their characteristics. Section 5 finally concludes the study.

2. Coordinated signal control problem

In this section, we formulate a *deterministic* coordinated signal control problem. After discussing the problem statement in Subsection 2.1, Subsections 2.2 and 2.3 describe the objective function and constraints, respectively. We then demonstrate the overall problem and make comparisons with existing formulations in Subsection 2.4.

2.1. Problem setting

Consider a two-way arterial corridor with M signalized intersections and $2 \times M$ crossroads, as shown in Fig.1. The set of intersections is denoted by $\mathcal{M} = \{1, 2, \ldots, M\}$, and the intersections are numbered $1, 2, \ldots, M$. The set of roads

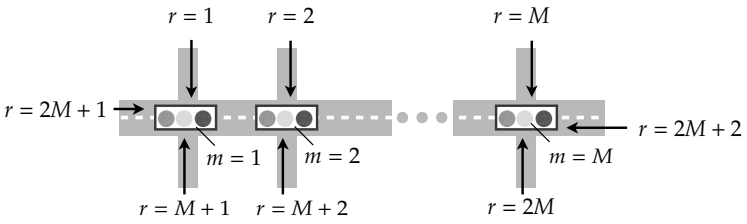


Fig. 1. Two-way arterial corridor with M signalized intersections and $2 \times M$ crossroads

is denoted by $\mathcal{R} = \{1, 2, \ldots, 2M + 2\}$; one and opposite directions of the arterial corridor (main road) are numbered $2M + 1$ and $2M + 2$, respectively. The crossroads are numbered according to the intersection being approached. Following Lo (2001), to simplify the exposition, turning is prohibited in the corridor. Thus, only two-phase signals are considered. Further, for notational simplicity the arterial corridor and crossroads are treated as being one-way only in this section and in Section 3, though the numerical experiments in Section 4 consider two-way traffic. The control time duration is $[0, I]$, and the target section of each road r is $[0, L^r]$.

We assume that each road is homogeneous and that the traffic flow on the road $r \in \mathcal{R}$ is characterized by a triangular fundamental diagram with the following parameters: u^r (forward wave speed), w^r (backward wave speed), q^r_{max} (saturation flow rate), and κ^r_j (jam density). The exogenous conditions required to describe the traffic dynamics on each road are the initial and the boundary conditions at the entrance and exit of the target section. The initial condition reflects the remaining queue before the target time duration[1]. The traffic demand that has arrived at the entrance of the road is given based on observations by traffic detectors[2].

Under the above conditions, our problem is to determine the *pretimed* signal setting that minimizes the total signal delay on all roads. This problem can be conceptually described as the following optimization problem:

$$\min_{\mathbf{s}} . \ \sum_{r \in \mathcal{R}} D^r(\mathbf{s}) \qquad \text{s.t.} \quad \mathbf{s} \in \mathcal{S}, \tag{1}$$

where $\mathbf{s}$ is an on-and-off signal control pattern vector with a 0–1 binary element representing the red or effective green phase at each discrete time unit Δt for each road, $D^r(\cdot)$ is the signal delay for each road under the given signal pattern and the initial and boundary conditions, and $\mathcal{S}$ is the feasible region of $\mathbf{s}$. It should be noted that the common signal control parameters (i.e., green splits and offsets) do not constitute the decision variables in this problem (1). This allows the modeling of the feasible region $\mathcal{S}$ to be significantly simplified (as discussed in Subsection 2.3 and 2.4), as the common signal control *parameters* are uniquely determined from the resulting signal control *pattern*.

2.2. Signal delay for a given signal control pattern

2.2.1. Variational theory of kinematic waves

For a given signal control pattern, we evaluate the signal delay of each road by using the Variational Theory (VT) (Daganzo, 2005a,b), which is a generalization of Newell's simplified kinematic wave theory (Newell, 1993). In this theory, the cumulative count of vehicles $N(t, x)$ at location $x \in [0, L]$ by time $t \in [0, T]$ is the state variable and is assumed to be differentiable (except along shock waves). The relationships among flow $q(t, x)$, density $\kappa(t, x)$, and cumulative flow $N(t, x)$ are described as $q(t, x) = \partial N(t, x)/\partial t$ and $\kappa(t, x) = -\partial N(t, x)/\partial x$. Then, the fundamental diagram $q(t, x) = Q(\kappa(t, x), x)$ can be expressed by the following Hamilton–Jacobi equation:

$$\partial N(t, x)/\partial t = Q(-\partial N(t, x)/\partial x, x). \tag{2}$$

In the VT, the problem (2) is solved by the variational method: the cumulative flow N_P at point $P = (t_P, x_P)$ in the time-space region is determined by the following continuum optimization problem (Daganzo, 2005a):

$$N_P = \min_{B \in \mathcal{B}} . \left[N_B + \int_{t_B}^{t_P} \max_{\kappa \in [0, \kappa_j]} . \left(Q(\kappa, x) - \kappa x'(t, x) \right) \mathrm{d}t \right], \qquad (t, x) \in \phi_{BP} \tag{3}$$

[1] A probe car data can be used for setting the initial condition (for example, see Mehran et al., 2012).

[2] In our setting, the vehicle arrivals at the intersections (except the most upstream one) of the main road is endogenously determined in the traffic flow model. Also, there are no inflows from crossroads to the main road and outflows from the main road to crossroads.

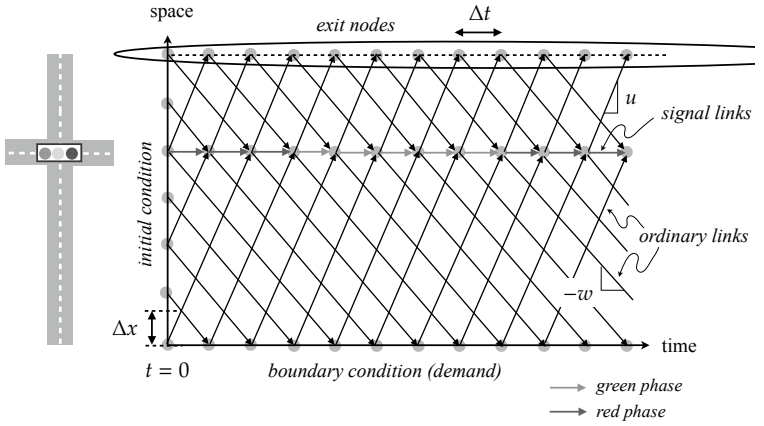


Fig. 2. Lopsided network on time-space region (VT network)

where ϕ_{BP} is the trajectory of the moving observer from the point B on the boundary $\mathcal{B}$ to the point P, $x'(t,x) \in [-w, u]$ is the velocity of the observer at $(t,x) \in \phi_{BP}$ and N_B is the known cumulative flow value at the point $B \in \mathcal{B}$ (i.e., initial and boundary conditions). The function, $\max_{k\in[0,\kappa_j]} . [Q(\kappa,x) - \kappa x'(t,x)]$, represents the maximum flow that can be experienced by the observer with velocity $x'(t,x)$ (*relative capacity*).

2.2.2. Solution method

The VT also provides a method for solving the continuum problem (3). This method is a (discrete) shortest path problem on a network consisting of nodes and links with the following two properties: (i) the slopes of links branching from each node which represent feasible wave speeds; (ii) link costs which represent the allowable change in the cumulative count of vehicles (i.e., relative capacity) along valid links (Daganzo, 2005b). We call this a *VT network*. For a homogeneous problem with a triangular fundamental diagram, the VT network becomes the lopsided network (*sufficient network*) shown in Fig. 2 wherein the mesh resembles the fundamental diagrams.

In this network, the nodes (gray circles in the figure) are placed on the positions of the boundaries of homogeneous road sections (i.e., the entrance and exit of the target road section and signalized intersections) and the initial time on a time-space region, and the set is denoted by $\mathcal{V}$. The links are classified into two types: *ordinary* and *signal* links. The ordinary link has a slope of u or $-w$, and its cost in unit time Δt is 0 or $q_{max}\Delta t$. The signal link has a zero slope, and its cost is $q_{\max}\Delta t$ if the link is an effective green phase and zero otherwise; this is represented by a binary variable $s \in \{0,1\}$: $q_{max}\Delta t \cdot s$. We denote the sets of ordinary links and signal links by $\mathcal{L}_o$ and $\mathcal{L}_s$, respectively. Each node is identified by an index $i = (t_i, x_i)$, and each link and its cost are denoted by a node pair (i, j) and c_{ij}.

As in the continuum problem (3), the cumulative flow N_i at node i is determined by solving the shortest path problem on the VT network with multiple origins (i.e., initial/boundary nodes) and a single destination (i.e., node i). Furthermore, if we add to the VT network a dummy node o and dummy links connecting the dummy node and the initial/boundary nodes (so that the link cost = the known cumulative flow value of each node), the cumulative flows at all the nodes can be calculated simultaneously as a shortest path problem with one origin (i.e., dummy node) and many destinations (i.e., all other nodes)[3]. Note that the dummy links are assumed to be included in the set $\mathcal{L}_o$.

2.2.3. Signal delay evaluation problem

For a given signal pattern $\mathbf{s}^r$, the signal delay of each road r is defined as the sum of the vertical differences between the virtual cumulative arrivals and the cumulative departures at the exit during $t \in [0, I]$. The virtual cumulative arrival represents the cumulative flow at the exit in the case in which vehicles are not delayed. It is obtained by shifting the cumulative arrivals at the entrance by time t so that it reaches the exit after the free flow travel time passing, i.e., $t + L/u$. Thus, the signal delay is written as follows (for notational brevity, we here omit road index r):

$$D(\mathbf{s}) = \sum_{k\in\mathcal{K}} N(k\Delta t - L/u, 0)\Delta t - \sum_{k\in\mathcal{K}} N(k\Delta t, L)\Delta t = U - \sum_{j\in\mathcal{V}_{exit}} N_j \Delta t, \tag{4}$$

[3] This shortest path problem can be solved very efficiently (i.e., in linear time: $O(|\mathcal{V}| + |\mathcal{L}|)$) because the VT network is a directed acyclic graph.

where $\mathcal{K} = \{0, 1, 2, \dots\}$ is the set of finite time intervals with unit time Δt (during $[0, I]$) and $\mathcal{V}_{exit}$ is the set of nodes at the exit of the VT network at each time interval $k \in \mathcal{K}$. The first and second terms of this equation represent the virtual cumulative arrivals and cumulative departures at the exit, respectively. Because the cumulative arrivals are known (denoted by U), only the cumulative departures need to be evaluated. Using the VT, this evaluation problem can be formulated as the following linear programming:

$$D(\mathbf{s}) = U - \min_{\mathbf{y} \geq \mathbf{0}} . \sum_{(i,j) \in \mathcal{L}_o \cap \mathcal{L}_s} c_{ij} y_{ij} \tag{5}$$

$$\text{s.t.} \quad \sum_i y_{ij} - \sum_i y_{ji} = \delta_{id} \quad \forall i \in \mathcal{V}, \tag{6}$$

or its dual formulation:

$$D(\mathbf{s}) = U - \max_{\mathbf{N}} . \sum_{j \in \mathcal{V}_{exit}} N_j \Delta t \tag{7}$$

$$\text{s.t.} \quad N_j \Delta t \leq N_i \Delta t + c_{ij} \quad \forall (i, j) \in \mathcal{L}_o \cap \mathcal{L}_s \tag{8}$$

where δ_{id} is Kronecker's delta (i.e., 1 if i is the destination node $d \in \mathcal{V}_{exit}$, zero otherwise). Note that, without loss of generality, $N_o = 0$.

Problems (5) and (7) are well-known linear programming formulations of the shortest path problem (Papadimitriou and Steiglitz, 1982). The primal formulation (5) is the problem of finding the shortest *paths* (i.e., wave or valid paths) to all the destinations and is a direct discrete version of the continuum problem (3). The dual formulation (7) is the problem of finding the *costs* of the shortest paths (i.e., cumulative flows). The dual formulation, which is mainly used in addressing our deterministic signal optimization problem, may be useful in traffic engineering applications because the cumulative flow quantity is the decision variable. However, the point of view given by the primal formulation (wave paths) is also important for our stochastic extension (see Section 3) and in other applications, such as the analytical approximation of a macroscopic fundamental diagram by the VT (e.g., Daganzo and Geroliminis, 2008; Leclercq and Geroliminis, 2013; Daganzo and Lehe, 2016).

Before concluding this subsection, we show the vector-matrix representations of the signal delay formulations that are used later in the paper. If $\mathbf{A}$ is the node-link incidence matrix of the VT network (including dummy links), $\mathbf{b}$ is the LHS of the flow conservation (6), $\mathbf{c} = \{c_{ij}\}_{(i,j) \in \mathcal{L}_o \cap \mathcal{L}_s}$ and $\mathbf{N} = \{N_i\}_{i \in \mathcal{V}}$. Then, the problems (5) and (7) can be expressed by

$$D(\mathbf{s}) = U - \min_{\mathbf{y} \geq \mathbf{0}} . \mathbf{c}^T \mathbf{y} \qquad \text{s.t.} \quad \mathbf{A}\mathbf{y} = \mathbf{b}, \tag{9}$$

$$D(\mathbf{s}) = U - \max_{\mathbf{N}} . \mathbf{b}^T \mathbf{N} \qquad \text{s.t.} \quad \mathbf{A}^T \mathbf{N} + \mathbf{c} \geq \mathbf{0}. \tag{10}$$

Without loss of generality, we here assume that $\Delta t = 1$.

2.3. Signal constraints

2.3.1. Minimal signal constraints and signal-constraint network

In this study we consider an on-and-off pattern (green-and-red phase sequence) that satisfies the minimal physical constraints to be the decision variable in the signal optimization problem. The minimal constraints for each intersection $m \in \mathcal{M}$ are set as follows:

- The signals of both the main road and a crossroad must not have an effective green phase at the same time.
- There is lost time when the signal switches.

An example of a feasible signal pattern is shown in the upper part of Fig. 3, which represents a pair of signal links (and nodes) extracted from the VT networks for both main and crossroads at the same intersection.

To describe these inter-relations between signal timings on the main and crossroads, we introduce a specially designed network, called a *signal-constraint (SC) network*, that comprises the signal links of the VT network for both main and crossroads as well as dummy links and dummy origins and destinations (blue nodes), as shown in the lower part of Fig. 3. In this network, the nodes at the mth intersection of the VT network of the main road and those of crossroad $r = m$ are connected by dummy links for ensuring lost times[4] (i.e., the time difference between two nodes

[4] This duration represents the sum of the clearance and start-up lost times (≈ the sum of yellow and all-red times).

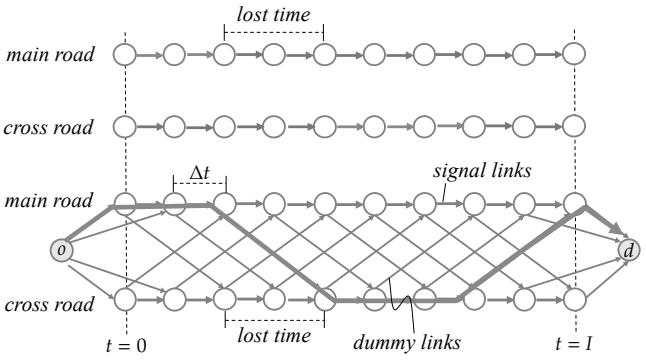


Fig. 3. Signal-constraint network (lost time = $2\Delta t$)

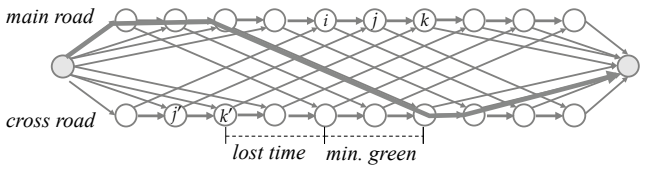


Fig. 4. Modified signal-constraint network (lost time = $2\Delta t$ and mim. green time = $2\Delta t$)

connected by the dummy link is the lost time: $2\Delta t$ in this case), and the dummy origin and destination are placed at the beginning/ending times. Any path between the origin and destination then satisfies the above mentioned signal constraints if the signal links are set to the effective green phase when the path (unit flow) passes through these links.

From the structure of the SC network, it is almost obvious that the minimal signal constraints reduce to a flow conservation like that in Eq. (6). Let $\overline{\mathcal{V}}^m$ be the set of nodes, $\overline{\mathcal{L}}^m$ be the set of links and z_{ij}^m be the binary flow for link (i, j) in the SC network. The minimal constraints for the mth intersection are then formulated as follows:

$$\sum_i z_{ij}^m - \sum_i z_{ji}^m = \delta_{id} \quad \forall i \in \overline{\mathcal{V}}^m \setminus \{o\}, \qquad z_{ij}^m \in \{0,1\} \quad \forall (i,j) \in \overline{\mathcal{L}}^m, \tag{11}$$

or equivalently,

$$\bar{\mathbf{A}}^m \mathbf{z}^m = \bar{\mathbf{b}}^m, \qquad \mathbf{z}^m \in \mathcal{Z}^m, \tag{12}$$

where $\bar{\mathbf{A}}^m$ is the node-link incidence matrix of the SC network, $\bar{\mathbf{b}}^m$ is the LHS of the Eq. (11), and $\mathcal{Z}^m \equiv \{0,1\}^{|\overline{\mathcal{L}}^m|}$ is the binary constraint on the vector $\mathbf{z}^m$. Note that, if the flows of signal links on the SC network for either the main or crossroads are set as binary variables, then the other variables automatically satisfy the binary constraint. Thus, the minimum number of binary variables required for each intersection is equal to the number of time intervals $|\mathcal{K}|$.

A path on the SC network $\mathbf{z}^m$ can be transformed linearly to the signal patterns of main and crossroads as there is a one-to-one correspondence between the signal links of the VT and SC networks (i.e., $z_{ij}^m = 1$ means that the corresponding link on the VT network is the effective green phase). Let $\mathbf{z}_s^m$ and $\mathbf{z}_d^m$ be the sub-vectors of $\mathbf{z}^m$ with respect to the signal links and dummy links, respectively. Then, the linear transformation can be expressed as

$$\left[\frac{\mathbf{s}_m^{2M+1}}{\mathbf{s}^m}\right] = \left[\begin{array}{c|c} \mathbf{I} & \mathbf{0} \end{array}\right] \left[\frac{\mathbf{z}_s^m}{\mathbf{z}_d^m}\right] \quad \Leftrightarrow \quad \bar{\mathbf{s}}^m = \mathbf{T}\mathbf{z}^m, \tag{13}$$

where $\bar{\mathbf{s}}^m \equiv [\ \mathbf{s}_m^{2M+1} \mid \mathbf{s}^m\ ]^T$ is the pair of signal pattern vectors of the main and crossroads for the mth intersection, $\mathbf{s}_m^{2M+1}$ is the signal pattern sub-vector of main road $r = 2M+1$ for the mth intersection, $\mathbf{T} \equiv [\ \mathbf{I} \mid \mathbf{0}\ ]$ is the transformation matrix, and $\mathbf{I}$ is the identity matrix.

2.3.2. Minimum green time constraint

The minimum green time constraint is basically required for pedestrian clearance and is critical at urban intersections. This constraint may be formulated in a number of ways. For example, Lin and Wang (2004) formulated the

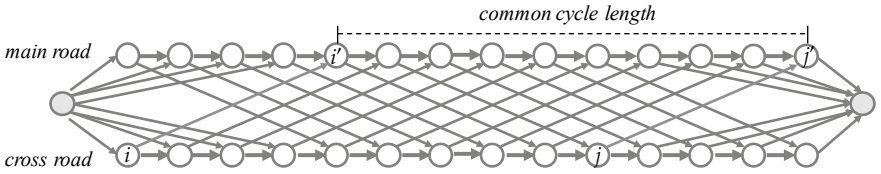


Fig. 5. Schematic of pretimed control constraint (lost time = $2\Delta t$, mim. green time = $2\Delta t$, and cycle length = $9\Delta t$)

constraint by using a moving time window concept, which directly constrains the green-and-red phase sequence (i.e., signal links) but requires a large set of linear constraints (i.e., three linear constraints are required for each signal link).

In this study, we adopt an alternative expression that is based on the SC network. This does not require any additional explicit constraints. Specifically, we modify the SC network as shown in Fig. 4. In the modified SC network (called a mSC network), the dummy link connects nodes of the main and crossroads so that the time difference between nodes is the sum of the lost times and minimum green times. Using this network, the signal constraints can still be formulated in the flow conservation form (12) but $\bar{\mathbf{A}}^m$ should be adapted to the mSC network (hereafter, we regard $\bar{\mathbf{A}}^m$ as the node-link incidence matrix of the mSC network).

The transformation matrix $\mathbf{T}$ should also be modified because the path on the mSC network does not pass through the signal links that represent the minimum green times. For example, let us consider signal link (i, j) in Fig. 4 and denote the phase of corresponding signal link of VT network by s_{ij}^{2M+1}. Then, s_{ij}^{2M+1} should be expressed by

$$s_{ij}^{2M+1} = z_{ij}^m + z_{j'j}^m + z_{k'k}^m.$$

The first term is equal to one if the path passes through the link (i, j), exactly as in the Eq. (13). However, if the path passes through either of the dummy links (j', j) or (k', k), s_{ij}^{2M+1} should still be equal to one, even though the path does not pass through link (i, j). This is ensured by the second or third terms. This example demonstrates that the transformation is still linear. Therefore, using the matrix $\mathbf{B}^m$, which describes the relationships between the signal links of the VT networks and dummy links of the mSC network, the transformation (13) is modified as follows:

$$\left[\frac{\mathbf{s}_m^{2M+1}}{\mathbf{s}^m}\right] = \left[\begin{array}{c|c} \mathbf{I} & \mathbf{B}^m \end{array}\right] \left[\frac{\mathbf{z}_s^m}{\mathbf{z}_d^m}\right] \quad \Leftrightarrow \quad \bar{\mathbf{s}}^m = \mathbf{T}^m \mathbf{z}^m, \tag{14}$$

where $\mathbf{T}^m \equiv [\, \mathbf{I} \mid \mathbf{B}^m \,]$.

2.3.3. Constraints for pretimed control and offsets

To obtain a cyclic (pretimed) signal control pattern and its corresponding common signal control parameters, we need to set a fixed common cycle length and impose a set of linear constraints on the path on the mSC network. Fig. 5 shows a schematic of this constraint; if a path passes [or does not pass] through dummy link (i, i'), the path should also pass [or should not pass] through the dummy link (j, j') that terminates to the node exactly one common cycle time later. This constraint is mathematically expressed as $z_{ii'}^m - z_{jj'}^m = 0$. Because it is also linear, using the matrix $\mathbf{C}^m$ that describes the above mentioned relationship, the pretimed control constraint is expressed as

$$\mathbf{C}^m \mathbf{z}_d^m = \mathbf{0}. \tag{15}$$

Note that the offsets can be calculated from the solution of the signal optimization problem (i.e., the paths on the mSC networks) with this constraint: these are given as the difference in the starting times of the first effective green phase on the main roads for the adjacent intersections. Fig. 6 shows two mSC networks for intersections m and $m+1$, where the blue [red] links represent the dummy links from the crossroad to the main road [from the main road to the crossroad] that the path passes through. The left and right figures show the cases in which the relative offsets are 50% and -50%, respectively. It can be seen that it is the time difference between the terminal nodes of the first blue links for the adjacent intersections that determines the offset. Furthermore, if we impose this cycle-length constraint on the dummy links from the crossroad to the main road only, the offset is fixed but the green splits can change cycle by

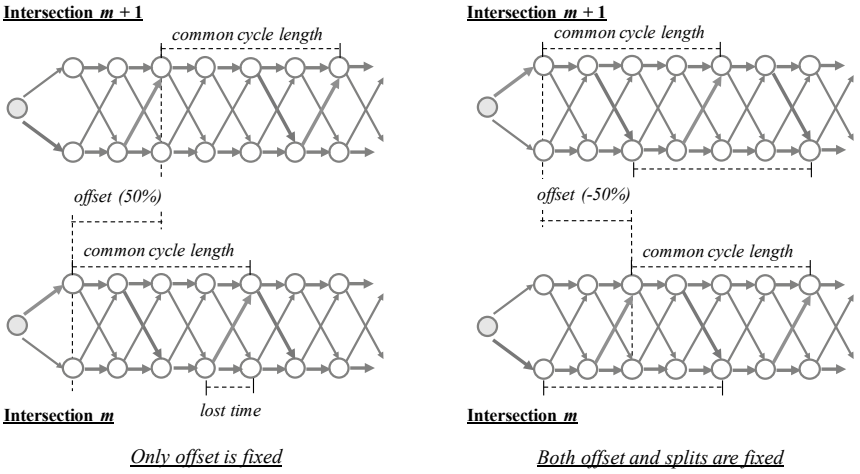


Fig. 6. Pretimed control constraint and offset (lost time = Δt, mim. green time = 0, and cycle length = $4\Delta t$). Left: offset = 50%; only offset is fixed. Right: offset = −50%; both offset and green splits are fixed.

cycle (see the left part of Fig. 6). This can be regarded as a semi-actuated coordinated signal control. In contrast, if we impose the constraint on the dummy links in both directions, both the offset and green splits are fixed cycle by cycle (see the right part of Fig. 6).

2.4. Overall problems

We next show the pretimed coordinated signal optimization problem by combining the objective function and signal constraints specified above. Because the objective function (9) or (10) is itself an optimization problem, a straightforward formulation of the overall problem leads to a *bilevel* problem:

- the upper level problem finds the paths on the SC networks for all intersections to minimize the total delay;
- the lower level problem determines the total delay by the VT with a signal pattern as a parameter.

Nevertheless, our problem can be formulated as a *single-level* problem[5] by using the dual formulation of delay (10) because the objective functions of both problems are described by exactly the same cumulative flows. That is,

$$\begin{aligned} & \min_{\mathbf{z}^m \in \mathcal{Z}^m} . \ \sum_{r\in\mathcal{R}} D^r(\mathbf{s}) \equiv \sum_{r\in\mathcal{R}} [U^r - \max_{\mathbf{N}^r} . \ (\mathbf{b}^r)^T \mathbf{N}^r] \\ \Rightarrow \quad & \max_{\mathbf{z}^m \in \mathcal{Z}^m,\ \mathbf{N}^r} . \ \sum_{r\in\mathcal{R}} (\mathbf{b}^r)^T \mathbf{N}^r \end{aligned} \tag{16}$$

$$\text{s.t.} \quad (\mathbf{A}_o^r)^T \mathbf{N}^r + \mathbf{c}_o^r \geq \mathbf{0} \qquad \forall r \in \mathcal{R} \tag{17}$$

$$(\mathbf{A}_s^r)^T \mathbf{N}^r + q_{max}^r \mathbf{s}^r \geq \mathbf{0} \qquad \forall r \in \mathcal{R} \tag{18}$$

$$\bar{\mathbf{A}}^m \mathbf{z}^m = \bar{\mathbf{b}}^m \qquad \forall m \in \mathcal{M} \tag{19}$$

$$\mathbf{C}^m \mathbf{z}_d^m = \mathbf{0} \qquad \forall m \in \mathcal{M} \tag{20}$$

$$\bar{\mathbf{s}}^m = \mathbf{T}^m \mathbf{z}^m \qquad \forall m \in \mathcal{M} \tag{21}$$

where the superscript r indicates the matrix/vector for road r, the matrices $\mathbf{A}_o^r$ and $\mathbf{A}_s^r$ are the sub node-link incidence matrices of the VT network with respect to the ordinary link set $\mathcal{L}_o^r$ and signal link set $\mathcal{L}_r^r$, respectively, $\mathbf{c}_o^r = \{c_{ij}^r\}_{(i,j)\in\mathcal{L}_o^r}$.

This is a mixed integer linear programming (MILP) that can be solved using commercial optimization software, such as Gurobi Optimizer and CPLEX. The objective function and the first and second constraints correspond to the total delay evaluation problem by the VT, the objective function and third and fourth constraints represent the

[5] The following fact is used: $\min . (-a) = -\max . a$.

problem of finding paths on the SC networks to minimize the total delay, the fifth constraint converts the paths on the SC networks into the signal patterns on the VT networks. This makes it challenging to treat the common cycle length as the decision variable. However, we can obtain an optimal cycle length, green splits and offsets simultaneously, if we optimize the problems for different common cycle lengths and choose the best one. Such a two-step optimization approach is reasonable because the cycle length optimization requires a one-dimensional search.

The proposed formulation has several advantages over the existing mathematical programming models for addressing the KW model-based *pretimed* signal setting problem (Lo, 1999, 2001; Lin and Wang, 2004). First, the lost times are correctly and naturally incorporated into the model by using the SC network concept, allowing examination of the optimal cycle length, which is the most important control parameter. Second, the size of the problem is quite small compared with that of Lo (2001). Specifically, since the traffic flow dynamics is formulated as a linear programming in the VT, it can be incorporated into the mathematical programming model *without introducing binary variables*. By contrast, the CTM-based formulation requires three binary variables for each cell for each time interval. The number of required constraints is also reduced substantially as the VT allows the link-based modeling[6]. Regarding the signal constraints, because the proposed ones are not based on the *if-then conditions*, the number of binary variables for each intersection is one-third that in Lo (2001) (i.e., the number of time intervals $|\mathcal{K}|$). This clearly shows that modeling of signal constraints is as important as modeling of traffic flow. Furthermore, it is inappropriate to further reduce the number of binary variables because it allows an unrealistic signal pattern (i.e., violates the first condition of the minimal signal constraints in Subsection 2.3.1). Thus, we can conclude that the number of required binary variables in the proposed MILP is minimized. Finally, the proposed formulation has a clear network structure. This may be useful when developing an exact/approximate solution algorithm and others. In fact, we employ the VT and SC network structures in the next section.

The following remark concludes this section. If we remove the cycle length constraint (20), the problem can produce a dynamic and very flexible signal pattern that could be used as a benchmark for evaluating actuated signal control logic. Furthermore, if we use the primal formulation of delay (9) as the objective function, the problem has an interesting structure and suggests a connection with an existing operations research problem. In particular, this problem is formulated as follows:

$$\max_{\mathbf{z}^m \in \mathcal{Z}^m} . \min_{\mathbf{y}^r \geq \mathbf{0}} . \sum_{r \in \mathcal{R}} \left[(\mathbf{c}_o^r)^T \mathbf{y}_o^r + (q_{max}^r \mathbf{s}^r)^T \mathbf{y}_s^r \right], \quad \text{s.t.} \quad \mathbf{A}^r \mathbf{y}^r = \mathbf{b}^r \quad \forall r \in \mathcal{R}, \quad \text{Eqs. (19)–(21)} \tag{22}$$

where $\mathbf{y}_o^r$ and $\mathbf{y}_s^r$ are link flow sub-vectors of each VT network with respect to the ordinary links and signal links, respectively. This problem can be interpreted as a *game between the road manager and waves*: the waves try to minimize their path costs on the VT network, while the road manager tries to choose the paths on the mSC network to maximize their (path) costs. This is a variant of the *shortest-path network interdiction* problem (Israeli and Wood, 2002), which is more generally known as the *attacker-defender model* (e.g., Bell et al., 2008). Because many studies have been conducted in this area, the known mathematical properties and solution methods may be applied to (22). However, exploring such possibilities is beyond the scope of this study and will be addressed in the future research.

3. Stochastic extension

This section extends the proposed *deterministic* formulation to a *stochastic* one. Specifically, we consider the stochastic fluctuations in traffic demand (boundary condition) such as Poisson arrivals at the entrance of a road section. This could largely affect the delay, particularly under the near saturation condition. Subsection 3.1 presents an exact expected delay evaluation problem as a stochastic shortest path problem. We also provide a method that narrows down the candidate shortest paths by using a property of the waves on the VT network. Subsection 3.2 provides an analytical approximation to the stochastic shortest path problem using the Clark's method (Clark, 1961). The accuracy and efficiency of this approximation method are validated in Subsection 3.3. Subsection 3.4 develops a solution method for the stochastic version of the signal optimization problem based on the Cross-Entropy (CE) method (Rubinstein and Kroese, 2004). For simplicity, this section assumes that each road is initially empty, although any traffic condition could be arranged.

[6] The similar VT-based traffic flow modeling (Mazaré et al., 2011) is recently used in the traffic signal control problems (Han et al., 2012; Li et al., 2014; Han et al., 2016a). However, they do not consider the modeling of realistic signal constraints discussed in Subsection 2.3.

3.1. Exact evaluation of the expected signal delay

Let the number of vehicles arriving at the entrance of each road section during time interval $t \in ((l-1)\Delta t,\ l\Delta t] - L/u$ be a random variable ϵ_l. Then, the cumulative demand by $t = k\Delta t - L/u$ is given by

$$N_k \equiv N(k\Delta t - L/u, 0) = \sum_{l=0}^{k} \epsilon_l, \tag{23}$$

assuming $N(-L/u, 0) = 0$ and $\epsilon_0 = 0$. It is assumed that the random variable ϵ_l follows a Poisson distribution with an average arrival rate parameter λ because the random variables should be non-negative and the distribution has been widely used in the traffic signal control literature. In this case, the cumulative flow N_i at each node i on the VT network and signal delay for each road D become random variables because cumulative demands at the boundary nodes are random variables. Hence, the expected signal delay of each road is expressed as follows:

$$\mathbb{E}[D] = \sum_{k \in \mathcal{K}} \mathbb{E}[N_k] - \sum_{j \in \mathcal{V}_{exit}} \mathbb{E}[N_j], \tag{24}$$

where we use the additivity of expected values. In this equation, the first term (i.e., virtual cumulative arrivals at the exit) is obtained easily by using the linearity of the expectation operator:

$$\mathbb{E}[N_k] = \mathbb{E}[\textstyle\sum_{l=0}^{k} \epsilon_l] = \lambda k \qquad \forall k \in \mathcal{K}. \tag{25}$$

Conversely, the second term (i.e., cumulative departures) is obtained by a stochastic shortest path problem. Specifically, according to Eq. (3), the expected cumulative departure at the exit at time $t = j\Delta t$ is given by

$$\mathbb{E}[N_j] = \mathbb{E}[\min_{k \in \{0,1,2,\ldots,j\}} .\{\textstyle\sum_{l=0}^{k} \epsilon_l + R_{k,j}\}] \qquad \forall j \in \mathcal{V}_{exit} \tag{26}$$

where $R_{k,j}$ is the cost along the trajectory ϕ_{kj} from the k-th boundary node at the entrance to the j-th exit node on the VT network[7]. Because each path cost comprises deterministic (relative capacity) and stochastic (cumulative demand) terms, we can regard this problem as a *probit-type* shortest path problem interestingly[8]. Based on this natural connection between a stochastic kinematic wave problem and the discrete choice theory, we develop a simple solution method for this problem (26).

One issue towards solving this problem is the existence of correlations among the path costs: the stochastic cumulative demand $\sum_{l=0}^{k} \epsilon_l$ at the k-th entrance node is positively correlated with the cumulative demands $\sum_{l=0}^{k'} \epsilon_l$ for all $k' < k$. Therefore, it is difficult to obtain an exact solution for the problem (26), and an approximation method is needed. Such a method is described in the next subsection. The other issue is that the number of candidate entrance nodes linearly increases with time. For example, if we consider $\Delta t = 1$ [sec] and evaluate the cumulative departure by time after fifteen minutes, the number of candidate entrance nodes becomes 900. This may be undesirable for both computational efficiency and accuracy of the approximation method. However, the number of candidate nodes can substantially decrease when the property of wave paths on the VT network is used. This fact is explained in the following.

Let us consider j-th exit node shown in the upper part of Fig. 7 as the evaluation node. The candidate entrance nodes initially are $k = \{0, 1, 2, \ldots, j\}$. The lower part of this figure shows the deterministic cost from each entrance node to j-th exit node. By the definition of each link cost, the deterministic cost from j-th entrance node to j-th exit node is zero; and it monotonically increases with decreasing of the number k. More specifically, if we compare the deterministic costs from the adjacent entrance nodes k and $k+1$ to the same exit node, the difference of their costs, $R_{k,j} - R_{k+1,j}$, must be 0 or $q_{max}\Delta t$: if the path from the node k meets the path from the node $k+1$ by passing through the "green phase" signal link, then the difference is $q_{max}\Delta t$ (Fig. 7 shows this case), zero otherwise.

Combining $R_{k,j} - R_{k+1,j} = 0$ with the monotonic increase of cumulative demand (i.e., $\epsilon_l \geq 0$), we have

$$\textstyle\sum_{l=0}^{k} \epsilon_l + R_{k,j} \leq \sum_{l=0}^{k+1} \epsilon_l + R_{k+1,j}. \tag{27}$$

[7] In our case, the initial condition can be ignored because the cost along trajectories from nodes at initial time $t = 0$ to exit nodes are always larger than others because we assume that each road is initially empty.

[8] To be exact, a Poisson distribution is different from a normal distribution. However, the normal distribution is a good approximation to the Poisson distribution if λk is relatively large.

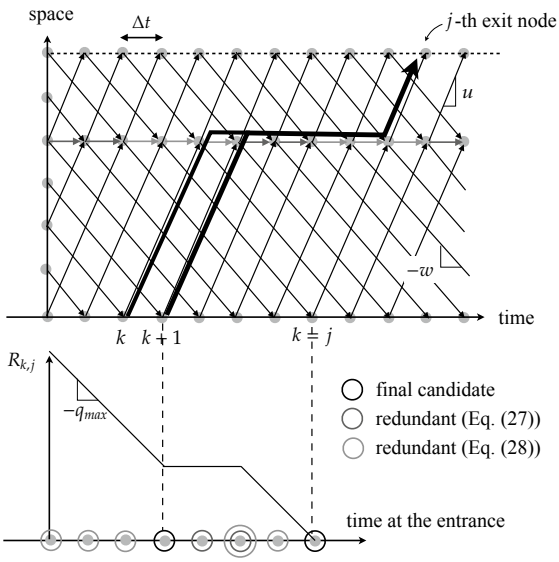


Fig. 7. Decision of the candidate node

This implies that $k+1$-th entrance node is redundant in the sense that it does not affect the expected cumulative departure value. Conversely, combining $R_{k,j} - R_{k+1,j} = q_{max}\Delta t$ with the consistency condition that the arrival flow rate during Δt is not greater than the maximum traffic flow rate defined by FD (i.e., $\epsilon_l \leq q_{max}\Delta t$), we obtain

$$\sum_{l=0}^{k} \epsilon_l + R_{k,j} \geq \sum_{l=0}^{k+1} \epsilon_l + R_{k+1,j}. \tag{28}$$

This means that k-th entrance node is redundant. By examining the conditions (27) and (28) for all initial candidate nodes and eliminating redundant nodes, we get the final candidate nodes which correspond to the valleys of the deterministic cost graph (in Fig. 7, $k+1$-th and j-th entrance nodes are the final candidates). Note that the present method is applicable for the case with non-empty initial conditions even though the shape of the deterministic cost graph changes (see Takayasu et al., 2016).

The reduction of the number of candidate nodes depends on the number of intersections and the offsets in the road section. For instance, if we consider a road with one intersection controlled by a pretimed signal, the number of candidate nodes is one per cycle. For an arterial road with several junctions, the number of candidate nodes depends on the offsets (i.e., the shape of the deterministic cost graph); however, the maximum number is equal to the number of junctions per cycle. Therefore, it is concluded that the number of candidate nodes is substantially reduced when a pretimed control with relatively long common cycle length is considered. Note that Eqs. (27) and (28) are not approximations but the exact conditions.

3.2. Analytical approximation of expected signal delay

In this subsection, problem (26) is analytically approximated. Clark's method (Clark, 1961) is used to evaluate the multiple integral of the correlated stochastic variables. This method is well known as the classical analytical approximation for a multinominal probit model (e.g., Daganzo et al., 1977). Furthermore, Deng et al. (2013) recently applied this method to a stochastic three detector problem[9].

Clark's method approximates the distribution of maximum values for two normal random variables as a new normal distribution. Let U_1, U_2, U_3 be the random variables that follow the multivariate normal distribution whose means, standard deviations and correlation coefficients are $\{V_1, V_2, V_3\}$, $\{\sigma_1, \sigma_2, \sigma_3\}$ and $\{\rho_{12}, \rho_{13}, \rho_{23}\}$. Then the first moment

[9] Because the three detector problem is a binary choice, they actually did not use the approximation but use an exact relation. Therefore, their purpose of using the Clark's method is different from ours.

v_1 and second moment v_2 of the random variable $\max(U_1, U_2)$ are exactly given by

$$v_1 = V_1\Phi(\gamma) + V_2\Phi(-\gamma) + a\phi(\gamma) \tag{29}$$

$$v_2 = (V_1^2 + \sigma_1^2)\Phi(\gamma) + (V_2^2 + \sigma_2^2)\Phi(-\gamma) + (V_1 + V_2)a\phi(\gamma) \tag{30}$$

$$V \equiv (V_1 - V_2)/a, \quad a^2 \equiv \sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2\rho_{12} \tag{31}$$

where $\phi(\gamma)$, and $\Phi(\gamma)$ are the probability density function and cumulative distribution function of the standard normal distribution, respectively. Using these formulas, the mean, variance, and correlation coefficient of $\max(U_1, U_2)$ can be expressed as

$$\mathbb{E}[\max(U_1, U_2)] = v_1 \tag{32}$$

$$\mathrm{Var}[\max(U_1, U_2)] = v_2 - v_1^2 \tag{33}$$

$$\rho[U_3, \max(U_1, U_2)] = \{\sigma_1\rho_{13}\Phi(\gamma) + \sigma_2\rho_{23}\Phi(-\gamma)\}/(v_2 - v_1^2)^{(1/2)}. \tag{34}$$

Clark (1961) further approximated the random variable $\max(U_1, U_2)$ as the following normal distribution:

$$\max(U_1, U_2) \sim \mathrm{N}(v_1, v_2 - v_1^2). \tag{35}$$

Repeating this approximation, we can get the approximate distribution of the maximum values for more than two normal random variables.

This method is applied to the problem (26). The final candidate entrance nodes are denoted sequentially as $(1, 2, 3, \ldots, h, \ldots, H)$ according to the increasing order of their original number k; the function relating the numbers h and k is indicated as ψ: $k = \psi(h)$. Then, the following problem is considered.

$$N_j = \min\left(\sum_{l=0}^{\psi(1)}\epsilon_l + R_{\psi(1),j}, ..., \sum_{l=0}^{\psi(H)}\epsilon_l + R_{\psi(H),j}\right) = -\max(U_1, ..., U_H) \tag{36}$$

where the element of the maximum function is $U_h = -\sum_{l=0}^{\psi(h)-1}\epsilon_l - R_{\psi(h),j}$ that is assumed to be normally distributed here. To approximate the Poisson arrivals by this random variable, we set the mean and variance of U_h as $V_h = -\lambda\psi(h) - R_{\psi(h),j}$ and $\sigma_h^2 = \lambda\psi(h)$, respectively. The correlation coefficient between different indices $h = h_1, h_2$ $(h_1 < h_2)$ are also given by $\rho_{h_1,h_2} = \sqrt{\psi(h_1)/\psi(h_2)}$.

Assuming $n_h \equiv -\max(U_1, ..., U_h)$, the approximating algorithm of the mean and variance of the cumulative departure at the exit is summarized as follows.

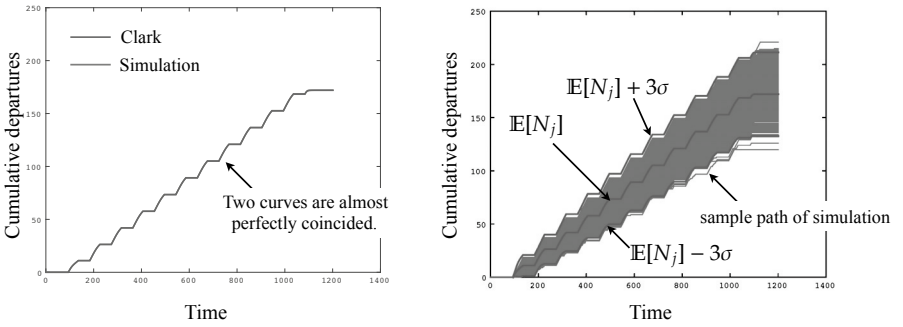
0. Set $h = 3$. Calculate $\mathbb{E}[-n_2]$, $\mathrm{Var}[-n_2]$, $\rho[U_h, \max(U_1, U_2)]$, $\forall h = 3, ..., H$ by Eqs. (32)–(34).
1. Using $\mathbb{E}[-n_{h-1}]$, $\mathrm{Var}[-n_{h-1}]$ and $\rho[U_h, \max(U_1, ..., U_{h-1})]$ calculated in the previous steps, calculate $\mathbb{E}[-n_h]$, $\mathrm{Var}[-n_h]$ by Eqs. (32) and (33).
2. If $h = H$ then go to Step 3; otherwise, calculate $\rho[U_{h'}, \max(U_1, \ldots, U_h)]$, $\forall h' = h + 1, ..., H$ by Eq. (34). Set $h := h + 1$ and go to Step 1.
3. Using the mean and variance of variable $-n_H$ calculated in the previous steps, the cumulative departure N_j is approximated by the following normal distribution: $N_j \sim \mathrm{N}(-\mathbb{E}[-n_H], \mathrm{Var}[-n_H])$, then stop.

3.3. Accuracy and efficiency of Clark's approximation

We examine the accuracy and efficiency of Clark's approximation for the problem (26) by a numerical experiment. Traffic flow on an arterial road with three signalized intersections is considered in this numerical experiment. The distances between intersections are 200 [m] (total length of the target road section is 800 [m]); the target time duration is $[0, 900]$ [sec]; and the discrete time unit is $\Delta t = 1$ [sec]. The parameters of triangular fundamental diagram of this road are set as $u = 40$ [km/h], $w = -15$ [km/h] and $q_{max} = 1600$ [veh/h]. All intersections are controlled by a pretimed signal with a common cycle length (60 or 90 [sec], here), and the green split of underlying road for each intersection is 50 [%]. For the offsets, we consider two patterns: one provides a "good" signal coordination, which implies that vehicles do not stop between intersections; the other provides a "bad" signal coordination, which implies

Table 1. Accuracy and efficiency of Clark's approximation

Cycle length [sec]	DoS [%]	Coordination	RMSE [veh]	Error [%]	Time (C) [sec]	Time (M) [sec]
90	85	good	0.17	0.19	0.60	50.1
90	85	bad	0.12	0.14	0.60	53.3
90	100	good	0.34	-0.25	0.62	51.8
90	115	good	0.32	-0.22	0.62	51.0
60	85	good	0.17	0.19	0.67	50.3



(a) Expected cumulative departures

(b) Variations of cumulative departures
(only 5,000 sample paths are shown)

Fig. 8. Mean and variations of cumulative departure curve (the first scenario)

that vehicles must stop between intersections. We set the average arrival flow rate, which is a Poisson distribution parameter, at the entrance such that the degree of saturation (DoS) becomes the target level ($85, 100, 115\%$, here).

In this experiment, we conduct $100,000$ times Monte Carlo simulations[10] of the VT for each scenario (see Table 1) and the mean value of simulated cumulative departures at each time at the exit is considered as the true value. As indices of the accuracy of Clark's approximation, we use the root mean square error (RMSE) of the mean value of cumulative departures and the relative error of the second term in Eq. (24):

$$\left\{\sum_{j\in\mathcal{V}_{exit}} \mathbb{E}[N_j] - \sum_{j\in\mathcal{V}_{exit}} \mathbb{E}_{Monte}[N_j]\right\} / \sum_{j\in\mathcal{V}_{exit}} \mathbb{E}_{Monte}[N_j] \tag{37}$$

where $\mathbb{E}_{Monte}[N_j]$ is the mean value of simulated cumulative departures. The second index is important because it represents the error of the expected signal delay (24), which is the objective function of the signal optimization problem. For coding and calculating the VT and the proposed approximation method, we used the MATLAB environment.

The results are summarized in Table 1; "Error [%]" indicates Eq. (37); "CPU time (C)" and "CPU time (M)" indicate the calculation times of the Clark's approximation and the Monte Carlo simulation of the VT, respectively. From this table, we can see that the accuracy of the Clark's approximation is significantly high for the different scenarios. This implies that the number of candidate entrance nodes, which depends on the cycle length and offsets, does not significantly affect its accuracy. However, the accuracy is slightly lower for near- and over- saturated cases. According to Horowitz et al. (1982), the Clark's approximation yields good results when the random components of the utilities (or costs) of different alternatives are independent or positively correlated moreover, they have variances that are not greatly different from one another. Both the conditions are satisfied in our case, thereby demonstrating good approximation of the cumulative flow. The mean and variations of cumulative departure curve for the first scenario is shown in Fig. 8. From these figures, we can see that the variance of cumulative departures as well as the mean value is well approximated by the proposed method. Moreover, the computational efficiency of the Clark's approximation is also high.

[10] In each simulation run, the number of vehicles arriving at the entrance of each road section during each unit time is generated by following a Poisson distribution (and the cumulative arrivals are calculated by Eq.(23)).

3.4. The Cross-Entropy method for stochastic coordinated signal control problem

Because the stochastic version of the coordinated signal control problem (i.e., the expected total delay minimization problem) is not reduced to a MILP, we develop a solution method based on the Cross-Entropy (CE) method. The CE method is universally applicable for estimating the probabilities of rare events and combinatorial optimization problems, and has many applications in various fields (e.g., Rubinstein and Kroese, 2004; de Boer et al., 2005). The CE method for combinatorial optimization problems comprises two steps: (i) generating the solution samples from a probability distribution; (ii) updating the probability distribution to generate a good solution sample with high probability (i.e., importance sampling). Contrary to other meta-heuristic methods, the CE method has almost no arbitrariness in the implementation, i.e., it basically does not require time-consuming tasks (e.g., appropriate settings of solution update rule, neighbors of a state, and other parameters). In addition, it can be naturally applied to network problems (e.g., traveling salesman problem) by introducing probability distributions on either the nodes or the links of the network. Therefore, the CE method is suitable for our problem in which the solution space is defined on the mSC network.

Here the solution method for the stochastic version of the coordinated signal control problem will be discussed. It is conceptually written as (i.e., the stochastic version of the problem (1))

$$\min_{\mathbf{z}\in\mathcal{Z}} . \mathbb{E}[D(\mathbf{z})] \equiv \sum_{r\in\mathcal{R}} \mathbb{E}[D^r(\mathbf{z})], \tag{38}$$

where $\mathbf{z}$ represents the paths on the mSC networks for all intersections and $\mathcal{Z}$ is its feasible region. We denote the optimal solution by $\mathbf{z}^*$ and the optimal value by $\gamma^* = \mathbb{E}[D(\mathbf{z}^*)]$.

In the CE method, a set of solution samples $\{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_\mathrm{N}\}$ (here N means the number of samples) are generated from a parametric probability distribution $f(\mathbf{z}; \mathbf{P})$ with a parameter $\mathbf{P}$. Because the solution of the problem (38) is given by the paths on the mSC networks, the probability distribution $f(\mathbf{z}; \mathbf{P})$ can be naturally described by

$$f(\mathbf{z}; \mathbf{P}) = \prod_{m\in\mathcal{M}} P^m(\mathbf{z}^m) \tag{39}$$

where $P^m(\mathbf{z}^m)$ is the probability of generating path $\mathbf{z}^m$ for intersection m. Considering the Markov chain from the origin to the destination on the mSC network for each intersection, the probability distribution (39) is further reduced to

$$f(\mathbf{z}; \mathbf{p}) = \prod_{m\in\mathcal{M}} \sum_{ij} p_{ij}^m \delta_{ij,\mathbf{z}^m} \tag{40}$$

where $\mathbf{p}$ is the parameter of the probability distribution that represents the transition probability matrices for all intersections, p_{ij}^m is the transition probability from node i to node j on the mSC network for intersection m, and $\delta_{ij,\mathbf{z}^m}$ is the element of the path-link incidence matrix for link (i, j): it is 1 if link (i, j) is on path $\mathbf{z}^m$ and zero otherwise.

However, if we choose a naive parameter $\mathbf{p}$, an impractical number of samples is required to estimate the global optimal solution for an enormous feasible region $\mathcal{Z}$. Thus, the CE method adopts an importance sampling. Let $g(\mathbf{z})$ be the *ideal* sampling function that generates the optimal solution with probability 1. It is given by:

$$g(\mathbf{z}) = \frac{I_{\{\mathbb{E}[D(\mathbf{z})]\le\gamma^*\}} f(\mathbf{z}; \mathbf{p})}{\sum_{\mathbf{z}\in\mathcal{Z}} I_{\{\mathbb{E}[D(\mathbf{z})]\le\gamma^*\}} f(\mathbf{z}; \mathbf{p})} \tag{41}$$

where $I_{\{\mathbb{E}[D(\mathbf{z})]\le\gamma^*\}}$ is an indicator function, taking value 1 if $\mathbb{E}[D(\mathbf{z})] \le \gamma^*$ and zero otherwise. Because the optimal value γ^* is unknown, our concern is to update the probability distribution $f(\mathbf{z}; \mathbf{P})$ so that it is "close" to $g(\mathbf{z})$ iteratively by using solution samples $\{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_\mathrm{N}\}$, where the distance between two distribution is measured by a cross-entropy (or Kullback-Leibler distance). According to de Boer et al. (2005), the problem of minimizing the cross-entropy for our problem is equivalent to the following problem:

$$\max_{\mathbf{p}} . \frac{1}{\mathrm{N}} \sum_{n=1}^{\mathrm{N}} I_{\{\mathbb{E}[D(\mathbf{z}_n)]\le\gamma\}} \ln \prod_{m\in\mathcal{M}} \sum_{ij} p_{ij}^m \delta_{ij,\mathbf{z}_n^m} \tag{42}$$

$$\text{s.t.} \quad \sum_i p_{ij}^m = 1 \qquad \forall i \in \overline{\mathcal{V}}^m, \ \forall m \in M \tag{43}$$

where γ ($\geq \gamma^*$) is a real number, which is called a"level", and the constraint is the normalization condition of the probability for each node on the mSC network. By solving this problem, a new parameter $\hat{\mathbf{p}}$ is achieved:

$$\hat{p}_{ij}^m = \frac{\sum_{n=1}^{\mathrm{N}} I_{\{\mathbb{E}[D(\mathbf{z}_n)]\leq\gamma\}}\delta_{ij,\mathbf{z}_n^m}}{\sum_{n=1}^{\mathrm{N}} I_{\{\mathbb{E}[D(\mathbf{z}_n)]\leq\gamma\}}\Delta_{i,\mathbf{z}_n^m}}, \tag{44}$$

where $\Delta_{i,\mathbf{z}_l^m}$ is the element of the path-node incidence matrix for node i: it is 1 if node i is on path $\mathbf{z}_n^m$ and zero otherwise. This update rule for the parameter $\mathbf{p}$ is very intuitive: among "good" solution samples whose expected delays are less than the level γ, the new probability of link (i, j) is given by the proportion of the number of passing counts of link (i, j) by the paths to that of node i by the paths.

It is important to set the level γ appropriately for obtaining a certain number of the "good" solution samples. The CE method thus updates the level γ as well as the parameter $\mathbf{p}$, iteratively. More specifically, the level γ for the next iteration is usually set as the best $100\rho\%$ value of the expected delays for the current solution samples. Here ρ is a parameter. Rubinstein and Kroese (2004) showed that any updating path of $(\mathbf{p}, \gamma)$ converges to $(\hat{\mathbf{p}}, \hat{\gamma})$, which can approximate $(\mathbf{p}^*, \gamma^*)$. The algorithm of the CE method can be summarized as follows.

0. Set the iteration count: $k = 1$. Choose sample size N and parameter ρ. Set the initial parameters $\mathbf{p}^{(1)}$ and $\gamma^{(1)}$.
1. Generate N samples $\{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_{\mathrm{N}}\}$ by the Markov chain on each mSC network with the transition probability $\mathbf{p}^{(k)}$. Calculate the expected total delay for each $\mathbf{z}_n$. Set $\gamma^{(k)}$ as the best $100\rho\%$ value.
2. Update the parameter $\mathbf{p}$ by Eq. (44), and denote it by $\mathbf{p}^{(k+1)}$.
3. If $\mathbf{p}^{(k+1)} = \mathbf{p}^{(k)}$ then stop; otherwise set $k := k + 1$ and go to Step 1.

Because each node on the mSC network basically has two outgoing links (signal and dummy links), the initial transition probabilities (i.e., parameter $\mathbf{p}^{(1)}$) for these links may be 0.5. In Step 1, by using the Markov chain for each intersection, we can efficiently generate all samples to be feasible. However, this is not always the case for problems with constraints that describe the relations between decision variables. It is not required for the Markov chain from the origin to destination to generate pretimed signal patterns, because the entire path is determined by the first round trip between nodes of main and crossroads on the mSC network (See Fig 6). In addition, in Step 1, we can efficiently calculate the expected total delay by the method developed in Subsection 3.2. Furthermore, we can apply the CE method to the deterministic signal optimization problem by replacing the calculation of expected delay with the VT (i.e., Eqs. (9) or (10)).

4. Numerical experiment

In this section, we describe the numerical experiment to demonstrate the usefulness of the proposed signal optimization problems in examining the optimal control parameters (i.e., cycle length, green splits and offsets) for coordinate signal controls. In particular, we compare the following signal setting strategies and discuss their differences:

- Strategy S: All signal parameters (the cycle length, green splits and offsets) are optimized to minimize the expected total delay under the stochastic formulation.
- Strategy D: With a given cycle length, the green splits and offsets are simultaneously optimized to minimize the deterministic total delay assuming the uniform arrival.
- Strategy P (practical signal setting): the common cycle length is first determined by Webster's formula (Webster, 1958); the allocation of green splits for each intersection is then determined to equalize the degree of saturations of all phases; the offsets are finally determined to minimize the deterministic total delay.

Consider a two-way arterial corridor with three signalized intersections and three crossroads as shown in Fig. 9. The distance between intersections $m = 1$ and $m = 2$ is 150 [m], and 250 [m] between intersections $m = 2$ and $m = 3$. The fundamental diagram is assumed to be the same for all roads, and its parameters' values are same as those in Subsection 3.3. We consider the two-way traffic in the arterial corridor ($r = 7, 8$) and the one-way traffic for in crossroads ($r = 1, 2, 3$). Because only through traffic is considered, two-phase signal is considered. The lost time and minimum green time are the same for all phases and all intersections, and their values are 3 [sec] (6 [sec] in

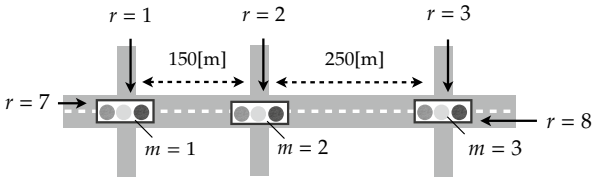


Fig. 9. A two-way arterial corridor with three signalized intersections and three crossroads

Table 2. Demand setting (average parameter of the Poisson and uniform arrivals)

Road No.	$r = 7$	$r = 8$	$r = 1$	$r = 2$	$r = 3$
Demand rate [veh/h]	700	700	150	700	400

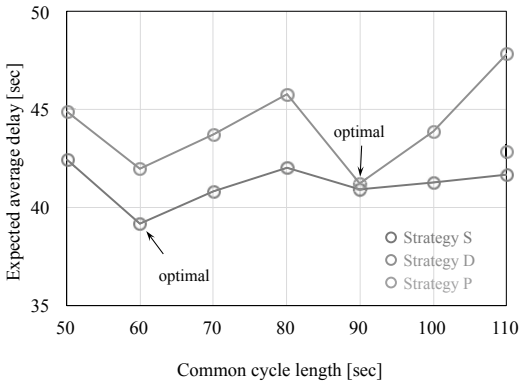


Fig. 10. Common cycle length versus expected average delays

each cycle) and 15 [sec], respectively. The demand rate (i.e., average parameter of the Poisson and uniform arrivals) for each road is summarized in Table 2. In this setting, intersection $m = 2$ is the critical intersection and the total flow ratio (= demand rate/saturation flow rate) is 0.875 (i.e., near saturated condition); the minimum and Webster's "optimal" cycle lengths for this intersection are 48 [sec] and 112 [sec], respectively. Therefore, we set the target range of the common cycle length is from 50 [sec] to 110 [sec]; within this range, we examine the cycle length while incrementing the common cycle length time by 10 [sec] (i.e., $50, 60, \ldots, 100, 110$ [sec]).

For the strategy S, we used the CE method (Subsection 3.4) to simultaneously optimize the green splits and offsets that minimize the expected total delay for each common cycle length. The parameters of the CE method are set as $(N, \rho) = (500, 0.1)$. For the strategy D, we used the Gurobi Optimizer to solve the MILP (16) that produces the green splits and offsets that minimize deterministic total delay for each common cycle length; the Clark's approximation is used for evaluating the expected total delay under the obtained green splits and offsets. For the strategy P, we first calculated the common cycle length and green splits for each intersection according to Webster's formulas; given the parameter values, the MILP (16) was solved to obtain the offsets that minimize the deterministic total delay[11].

Now, let us discuss the results. Fig. 10 shows the expected average delay per vehicle of all roads for each strategy and common cycle length. From this figure, we can see that the expected average delay under the strategy S is the lowest for all the common cycle lengths. This is because the strategy is the expected delay minimization strategy. Compared to the strategy S, the expected average delay under the strategy D is 7.3% higher on average, although those values are similar for 90 [sec] case. Moreover, for the 110 [sec] case, the performance of the strategy D is

[11] If we input the lost times and minimum green times to the proposed signal optimization problems such that the sum of these for each intersection equals the common cycle length, the green time of each phase becomes equal to its minimum green time. Therefore, for given common cycle length and green splits, only the offsets can be optimized.

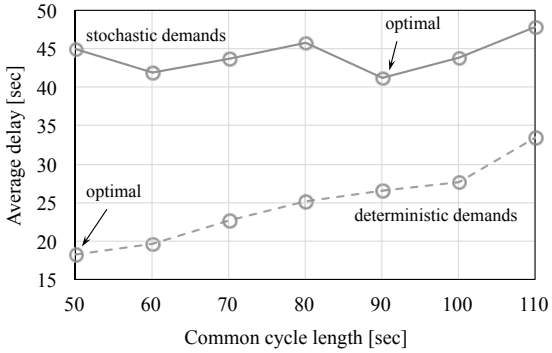


Fig. 11. Expected average delay versus deterministic average delay under the strategy D

Table 3. Splits and offsets under the selected cycle length (strategy S = 60 [sec], strategy D = 90 [sec], and strategy P = 110 [sec])

Strategy	Splits (main road) [%]			Relative offsets [%]	
	$m = 1$	$m = 2$	$m = 3$	1 – 2	2 – 3
S	72	52	59	95	55
D	82	52	69	16	94
P	83	50	63	13	96

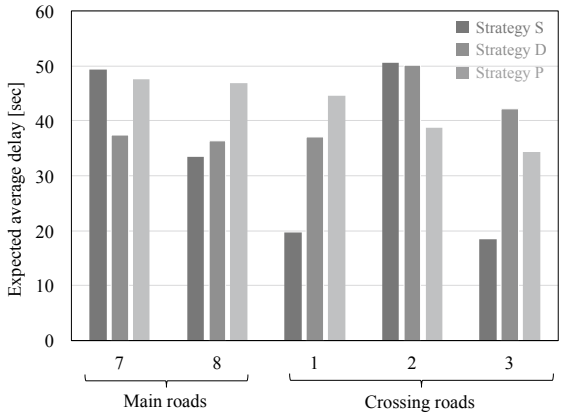


Fig. 12. Expected average delay for each road

much poor than that of the strategy P. These facts imply that the simultaneous optimization of control parameters in a *deterministic fashion* does not work well unless the cycle length is carefully chosen and is vulnerable to the fluctuation of traffic demand in some cases. The comparison between the expected and deterministic (i.e., uniform) average delays under the strategy D also supports this fact (see Fig. 11).

For the cycle length, strategy S selects the shortest length. This is much shorter than those of other strategies. The reasons are as follows. The strategy S can consistently consider the following properties for coordinated signal controls: (i) the demand fluctuation in the main road is reduced owing to the rectification through upstream signals; (ii) high dependence of the performance of signal coordination on the cycle length (e.g., Koshi, 1989; Jin and Yu, 2015) (in this experiment, 60 and 90 [sec] cases may lead to good signal coordination). As the result of combination of these two properties, the expected delay is not monotonically increasing with increasing of the cycle length. Because the strategy D can consider the latter property only, 90 [sec] is selected. In contrast, the strategy P (or Webster's formula) cannot consider these properties in principle, and thus the cycle length would become longer than necessary.

Next, we look into the green splits and (relative) offsets under the selected cycle length (see Table 3) to understand how the strategy S obtains the lower delay by using smaller available capacity (i.e., shorter cycle length). We also show the expected average delay for each road in Fig. 12 to understand the differences of the green splits and offsets

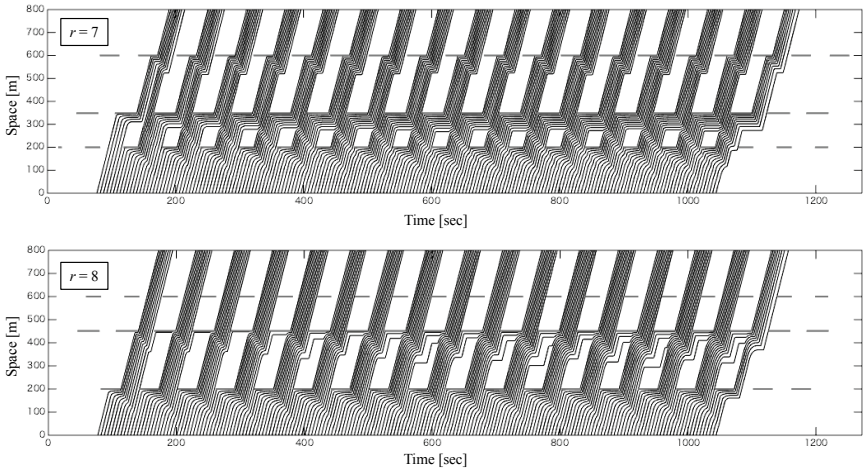


Fig. 13. Expected vehicle trajectories (Strategy S)

among different strategies. From the table and figure, we make the following observations[12]: (i) the strategy S gives more green splits to the minor crossroads $r = 1, 3$ than the strategies D and P; (ii) the strategy S sets the offsets to prioritize one direction (i.e., road $r = 8$), while the strategies D and P give the both directions equal priority; (iii) the variations of expected average delays among roads under the strategy S are higher than those under strategies D and P. One possible reason for the observation (i) is that the green splits of main road are sufficient to realize good signal coordination. The fact that the crossroads are directly affected by the demand fluctuations (i.e., Poisson arrivals) unlike the main road could also be the reason for observation (i). Observation (ii) may demonstrate the strong interdependency between the offsets and cycle length as mentioned earlier (because the cycle length in strategy S is much shorter than those in other strategies). By combining observations (i)–(iii), we may conclude that the strategy S utilizes the available capacity effectively by prioritizing roads based on the potential reduction of delay.

To confirm the validity of the above discussions, we show the expected vehicle trajectories (i.e., a contour on a three dimensional surface of expected cumulative curves) for main roads in Fig. 13–15. These figures were made by applying the Clark approximation to the problem (26) for every grid points in the VT network[13]. Interestingly, unlike the (deterministic) KW model, the instantaneous deceleration [acceleration] at the back [front] of queue is not always observed because the queue length fluctuates due to the random arrivals. From these figures, we can see that the effective green times are almost fully utilized under Strategy S, but those of the most downstream intersection are somewhat wasted under Strategies D and P. The observation (ii) can be also confirmed by looking at the expected vehicle trajectories. Furthermore, we can extract the important information of the queue evolution from these figures. For example, we can see that, for all strategies, the expected queue length at each intersection does not exceed the link length; but, the shockwave from the middle intersection $m = 2$ affects the upstream intersection $m = 1$ in road $r = 7$.

Finally, we show the distribution of the expected total delays under the selected cycle length in Fig. 16. From this figure, we can see that strategy S has the highest variance among strategies despite having the lowest expected delay. This implies that smaller available capacity is more susceptible to demand fluctuations. However, this figure also shows the interesting result: the frequency of the occurrence of high expected delay under strategy S is always less than that under other strategies. If this is true (for many cases), the strategy S can be regarded as a robust strategy (may be in a minimax sense) as well as the expected delay minimization strategy.

5. Conclusion

In this study, we considered a coordinated traffic signal control under both deterministic and stochastic demands. We first proposed a deterministic coordinated signal control problem as a MILP based on the variational theory with

[12] Although the similarity between the strategies D and P in the green splits and offsets is observed, this is not always the case because the strategy D clearly has more flexibility in determining these parameters.

[13] Even though the accuracy of the Clark's method is high for the problem (26), the small approximation errors are involved in these figures.

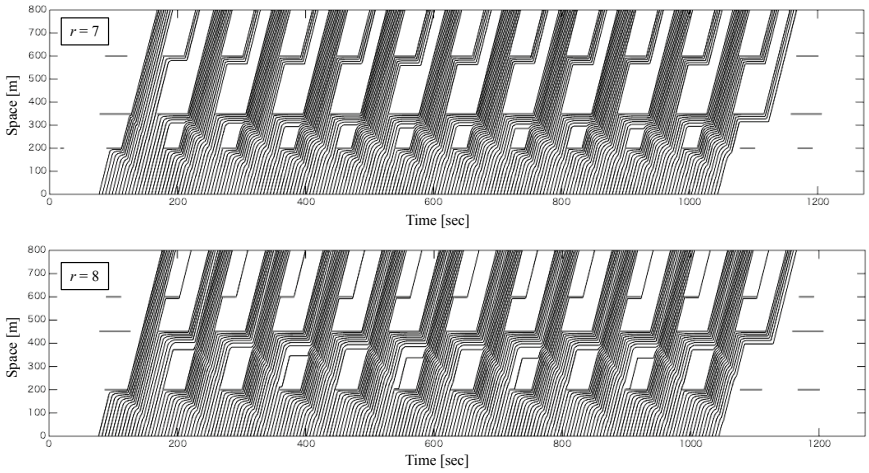


Fig. 14. Expected vehicle trajectories (Strategy D)

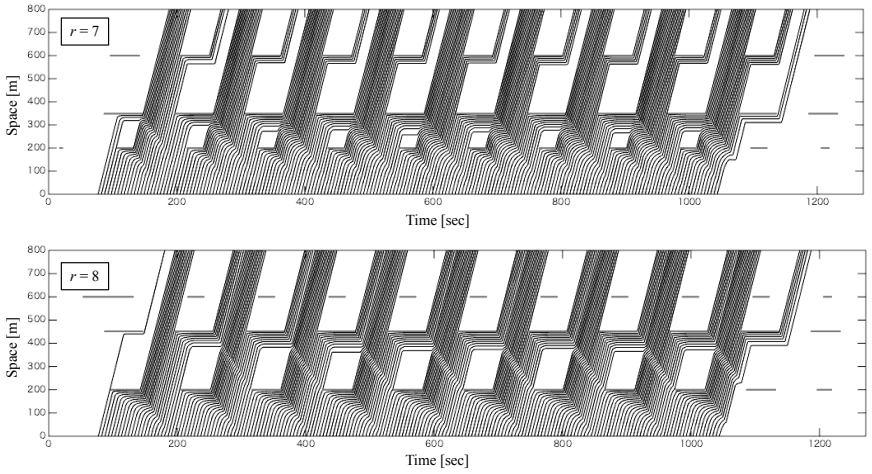


Fig. 15. Expected vehicle trajectories (Strategy P)

signal parameter constraints by introducing the VT network and the signal constraint (SC) network. The resulting MILP has a clear network structure and requires fewer binary variables and constraints as compared with those in the existing formulations; thus, it can be solved efficiently. Using this network structure, we then successfully extended the problem so as to treat the stochastic fluctuations in traffic demands. The expected delay was approximately evaluated by applying the Clark's method to a stochastic shortest path problem on the VT network. For the solution method, we employed a cross-entropy method for the stochastic version of the coordinated traffic signal control using the SC network. We finally examined the optimal signal control parameters for coordinated signal control under both the deterministic and stochastic demands and discussed their characteristics.

Although we investigated the effectiveness of the expected delay minimization control, the optimal signal parameters of the control depends on various factors, including the demand level, the number of intersections, and the distance of intersections. Therefore, a systematic numerical experiment should be conducted to explore the general property of the control. It would be also valuable to examine the performance of the control for the oversaturated conditions and time-dependent demands.

To consider more realistic situations such as turning movements (e.g., Lo et al., 2001; Han et al., 2016b), lane configurations (e.g., Li, 2011) and multiple phasing plans, we need to modify the proposed methods. For example, to consider the turning movements, appropriate node models (e.g., Tampère et al., 2011; Jabari, 2016) need to be incorporated into the proposed models. This makes the proposed problem more complicated because additional

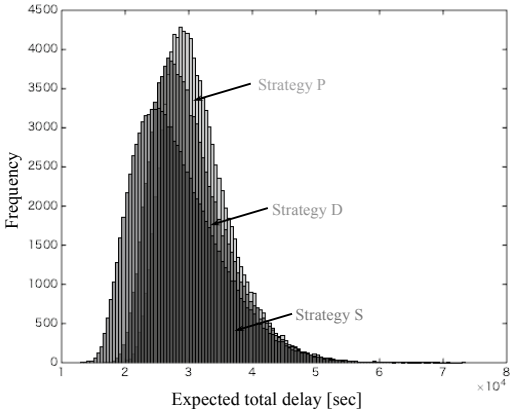


Fig. 16. Frequency distribution of total expected delays (100,000 samples)

binary variables should be introduced into the problem (Lo et al., 2001; Han et al., 2016b). However, the advantages of the proposed problem will remain unchanged because such an extension also makes the CTM-based problems even more complicated. On the other hand, it is easy to extend the SC network to the case with a multiple phasing plan if the phase sequence is pre-determined. Specifically, it is enough to connect the signal links for each phase with those for the next phase by dummy links for ensuring lost times, which enables us to represent a signal pattern as a path on the SC network. Then, the linearity of constraints remain unchanged, but we need to modify the transformation matrix $\mathbf{T}$ as in the case in Subsection 2.3.2. Note that these modifications will not change our basic framework.

From the viewpoint of the traffic flow theory, the simple analytical approximation method of expected delays (Subsection 3.2) is regarded as being a stochastic kinematic wave model based on the variational theory. Indeed, we can easily consider the fluctuation of the fundamental diagram as well as the stochastic boundary conditions (see Takayasu et al., 2016). By contrast, the different stochastic kinematic wave models based on the CTM were proposed (Sumalee et al., 2011; Jabari and Liu, 2012). Therefore, the comparison of these models with the proposed model will be an interesting topic for future work.

Acknowledgements

The authors express their gratitude to four anonymous referees for their careful reading of the manuscript and useful suggestions. This research is partially supported by JSPS KAKENHI Grants (26220906, 15H02270, 16K18163).

References

Bell, M.G.H., Kanturska, U., Schmöcker, J.D., Fonzone, A., 2008. Attacker–defender models and road network vulnerability. Philosophical transactions. Series A 366, 1893–906.

Cantarella, G.E., de Luca, S., Di Pace, R., Memoli, S., 2015. Network signal setting design: Meta-heuristic optimisation methods. Transportation Research Part C 55, 24–45.

Cheng, C., Du, Y., Sun, L., Ji, Y., 2016. Review on theoretical delay estimation model for signalized intersections. Transport Reviews 36, 479–499.

Clark, C.E., 1961. The greatest of a finite set of random variables. Operations Research 9, 145–162.

Daganzo, C.F., 2005a. A variational formulation of kinematic waves: basic theory and complex boundary conditions. Transportation Research Part B 39, 187–196.

Daganzo, C.F., 2005b. A variational formulation of kinematic waves: Solution methods. Transportation Research Part B 39, 934–950.

Daganzo, C.F., Bouthelier, F., Sheffi, Y., 1977. Multinomial probit and qualitative choice: A computationally efficient algorithm. Transportation Science 11, 338–358.

Daganzo, C.F., Geroliminis, N., 2008. An analytical approximation for the macroscopic fundamental diagram of urban traffic. Transportation Research Part B 42, 771–781.

Daganzo, C.F., Lehe, L.J., 2016. Traffic flow on signalized streets. Transportation Research Part B 90, 56–69.

de Boer, P.T., Kroese, D.P., Mannor, S., Rubinstein, R.Y., 2005. A tutorial on the cross-entropy method. Annals of Operations Research 134, 19–67.

Deng, W., Lei, H., Zhou, X., 2013. Traffic state estimation and uncertainty quantification based on heterogeneous data sources: A three detector approach. Transportation Research Part B 57, 132–157.

Han, K., Friesz, T.L., Yao, T., 2012. A link-based mixed integer LP approach for adaptive traffic signal control. arXiv:1211.4625.

Han, K., Liu, H., Gayah, V.V., Friesz, T.L., Yao, T., 2016a. A robust optimization approach for dynamic traffic signal control with emission considerations. Transportation Research Part C 70, 3–26.

Han, K., Piccoli, B., Szeto, W., 2016b. Continuous-time link-based kinematic wave model: formulation, solution existence, and well-posedness. Transportmetrica B 4, 187–222.

Horowitz, J.L., Sparmann, J.M., Daganzo, C.F., 1982. An investigation of the accuracy of the Clark approximation for the multinomial probit model. Transportation Science 16, 382–401.

Israeli, E., Wood, R.K., 2002. Shortest-path network interdiction. Networks 40, 97–111.

Jabari, S.E., 2016. Node modeling for congested urban road networks. Transportation Research Part B 91, 229–249.

Jabari, S.E., Liu, H.X., 2012. A stochastic model of traffic flow: Theoretical foundations. Transportation Research Part B 46, 156–174.

Jin, W.L., Yu, Y., 2015. Performance analysis and signal design for a stationary signalized ring road. arXiv:1510.01216.

Koshi, M., 1975. Cycle time control for coordinated traffic signals. Proceedings of the Japan Society of Civil Engineers 241, 125–133.

Koshi, M., 1989. Cycle time optimization in traffic signal coordination. Transportation Research Part A 23, 29–34.

Leclercq, L., Geroliminis, N., 2013. Estimating MFDs in simple networks with route choice. Transportation Research Part B 57, 468–484.

Li, Y., Canepa, E., Claudel, C., 2014. Optimal control of scalar conservation laws using linear/quadratic programming: Application to transportation networks. IEEE Transactions on Control of Network Systems 1, 28–39.

Li, Z., 2011. Modeling arterial signal optimization with enhanced cell transmission formulations. Journal of Transportation Engineering 137, 445–454.

Lin, W.H., Wang, C., 2004. An enhanced 0-1 mixed-integer LP formulation for traffic signal control. IEEE Transactions on Intelligent Transportation Systems 5, 238–245.

Lo, H.K., 1999. A novel traffic signal control formulation. Transportation Research Part A 33, 433–448.

Lo, H.K., 2001. A cell-based traffic control formulation: Strategies and benefits of dynamic timing plans. Transportation Science 35, 148–164.

Lo, H.K., Chang, E., Chan, Y.C., 2001. Dynamic network traffic control. Transportation Research Part A 35, 721–744.

Maher, M., Liu, R., Ngoduy, D., 2013. Signal optimisation using the cross entropy method. Transportation Research Part C 27, 76–88.

Mazaré, P.E., Dehwah, A.H., Claudel, C.G., Bayen, A.M., 2011. Analytical and grid-free solutions to the Lighthill-Whitham-Richards traffic flow model. Transportation Research Part B 45, 1727–1748.

Mehran, B., Kuwahara, M., Naznin, F., 2012. Implementing kinematic wave theory to reconstruct vehicle trajectories from fixed and probe sensor data. Transportation Research Part C 20, 144–163.

Newell, G.F., 1993. A simplified theory of kinematic waves in highway traffic, part I: General theory. Transportation Research Part B 27, 281–287.

Papadimitriou, C.H., Steiglitz, K., 1982. Combinatorial Optimization: Algorithms and Complexity. Prentice-Hall.

Papageorgiou, M., Kiakaki, C., Dinopoulou, V., Kotsialos, A., 2003. Review of road traffic control strategies. Proceedings of the IEEE 91, 2043–2067.

Park, B.B., Messer, C.J., Urbanik II, T., 2000. Enhanced genetic algorithm for signal-timing optimization of oversaturated intersections. Transportation Research Record 1727, 32–41.

Rubinstein, R.Y., Kroese, D.P., 2004. The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning. Springer.

Sumalee, A., Zhong, R., Pan, T., Szeto, W., 2011. Stochastic cell transmission model (SCTM): A stochastic dynamic traffic model for traffic state surveillance and assignment. Transportation Research Part B 45, 507–533.

Takayasu, A., Hara, Y., Wada, K., Kuwahara, M., 2016. Traffic state estimation considering stochasticity of input data based on variational theory. Presented at the 21st International Conference of Hong Kong Society for Transportation Studies.

Tampère, C.M., Corthout, R., Cattrysse, D., Immers, L.H., 2011. A generic class of first order node models for dynamic macroscopic simulation of traffic flows. Transportation Research Part B 45, 289–309.

Webster, F.V., 1958. Traffic signal settings. Road Research Technical Paper No. 39 .